

HOW INFORMATION TOOLS WORKING GROUP OCR TEST DOCUMENTS

Document #	Quality	Pages	Content	Summary Useful? (no OCR possible/type in)
DOTS-GENERATED-552	readable	1 page memo	request for reports	no
DOTS-GENERATED-553	readable	1 page memo	who to inform in case of disaster	no
DOTS-GENERATED-554	readable	1 page memo	river conference	no
DOTS-GENERATED-554	readable	1 page memo	disaster planning	no
DOTS-GENERATED-555	barely	1 page cover	??	no
DOTS-GENERATED-557	readable	2 page letter	conference cooperation	no
DOTS-GENERATED-559	readable	4 page letter	hanford works survey response	no
DOTS-GENERATED-563	barely	3 page letter	monte conference bickering	no
DUN-460	readable	15 monthly report	contamination control - technical	yes
DUN-615	readable	15 monthly report	contamination control - technical	yes
DUN-772	readable	17 monthly report	contamination control - technical	yes
DUN-914	readable	16 monthly report	contamination control - technical	yes
DUN-1008	readable	15 monthly report	contamination control - technical	yes
DUN-1139	readable	15 monthly report	contamination control - technical	yes
DUN-1381	readable	15 monthly report	contamination control - technical	yes
DUN-1585	readable	13 monthly report	contamination control - technical	yes
DUN-1631	readable	14 monthly report	contamination control - technical	yes
DUN-1775	readable	14 monthly report	contamination control - technical	yes
DUN-1854	readable	15 monthly report	contamination control - technical	yes
DUN-1248	readable	15 monthly report	contamination control - technical	yes
DUN-1936	readable	15 monthly report	contamination control - technical	yes
DUN-2029	readable	15 monthly report	contamination control - technical	yes
DUN-2243	readable	15 monthly report	contamination control - technical	yes
DUN-2273	readable	27 monthly report	contamination control - technical	yes
DUN-2361	readable	17 monthly report	contamination control - technical	yes
DUN-2498	readable	16 monthly report	contamination control - technical	yes
DUN-2942	readable	19 monthly report	contamination control - technical	yes
DUN-3368	readable	19 monthly report	contamination control - technical	yes
DUN-3032	readable	8 annual report	chemicals discharged in Columbia River	yes
DUN-3935	readable	17 quarterly report	contamination control - technical	yes
DUN-4477	readable	16 quarterly report	contamination control - technical	yes
DUN-4847	readable	15 quarterly report	contamination control - technical	yes
DUN-5030	readable	12 quarterly report	contamination control - technical	yes
DUN-4104	readable	26 report	reactor operations/Columbia River	yes
DUN-5990-RD	barely	78 handwritten observ	??	no
DUN-5563	readable	11 quarterly report	contamination control - technical	yes
DUN-6706	readable	9 report	chemicals discharged in Columbia River	yes
DUN-WSU-3	readable	74 report	thermal hydraulic pilot model/Columbia	yes
DUN-WSU-8	readable	47 report part?	thermal hydraulic pilot model/Columbia	yes
HQ-1	readable	36 report	Columbia River symposium	yes
HW-11344	less	11 study	Columbia River fish radioactivity levels	yes
HW-11662	barely	14 report	Limnological radioactivity - technical	yes
HW-11861	barely	2 report	Dead Whitefish in Columbia River	yes
HW-12079	barely	11 report	radioactivity in belin ducks/Columbia	yes
HW-12573	barely	5 report	radioactivity in lower Columbia	yes
HW-12974	barely	5 proposal	radioecological study of Columbia	yes
HW-12975	barely	8 proposal supplement	supplement to above	yes
HW-15861	barely	15 meeting minutes	advisory group/Columbia River	yes
HW-17595	barely	41 meeting minutes	next year (1953) as above	yes
HW-17732	readable	50 meeting minutes	species meeting for river survey	yes
HW-20055	readable	66 report	aquatic invertebrates/irradiation effects	yes
HW-20958	readable	22 report	plankton radioactivity	yes
HW-21074	barely	8 report	review of studies program	yes
HW-21743	barely	7 letter	effects of breaching dam	yes
HW-22181	readable	39 minutes	Columbia River symposium	yes
HW-22851	readable	53 report	dispersion of radioactive wastes	yes
HW-32093	readable	53 report	radioactivity in fish in river	yes
HW-23797	barely	13 report	pile explosion effect on aquatic life	yes
HW-23793	barely	13 report	radiobiological problems in river	yes
HW-24356	barely	8 report	permissible limits of release into river	yes
HW-25167	barely	3 report	operating standards on river release	yes
HW-25191	readable	15 report	radioactivity in crustacea	yes
HW-28467	barely	16 report	Columbia River water as pile coolant	yes
HW-28778	readable	6 program	for testing above river water use/coolant	yes
HW-29217	less	21 report	temperature/radioactivity in Columbia gravel bed	yes

Document

DOTS-GENERATED - Documents from DOE-Richland's Declassified Document Tracking System
 DUN - Douglas United Nuclear (if further subcontracting was involved i.e. DUN-WSU-3 the subcontracting agency is placed second)
 HQ - Hanford Operations Office HW - Hanford Works/General Electric RL - Richland Office with multiple subcontractors

Quality

Readable - Words are not OCRable but can be read with minimal difficulty. The print quality is generally at minimal deterioration.
 Less - Less readable and not all OCRable, type quality may have smudges, characters are not easily decipherable.
 Barely - Considerable state of deterioration, difficult to read when printed out. Consistent light characters and/or smudges.

Pages and Content

The type of document and the number of pages. Some documents are time series. The subject may not necessarily match the content. There is no way to determine if a document is technical or just observations, etc.

APPENDIX 10. INFORMATION TOOLS OCR TEST DOCUMENTS AND KEY